Online Appendix

The Version with No Scale Effect

The contents are exactly the same as those of the original version of the paper until equation (8).

1 Baseline Model with No Patent Breadth Policy

We first build and solve our baseline model of cumulative innovation and growth, in order to familiarize readers with the key features of the setup. Accordingly, in this section, innovators will not face a binding patent breadth for their ideas to be both patentable and non-infringing (i.e., marketable). For now, we shall also assume that patents do not expire, namely that the patent length is infinite, so that incumbent patent-holders lose their monopoly power only when superseded by a new innovation. This will place the focus on the innovation process in our model, which we proceed to describe next. We will then close the model in general equilibrium and discuss its properties.

1.1 Model setup: The innovation process

Consider an economy composed of one industry, in which a continuum of differentiated varieties indexed by $j \in [0, 1]$ is produced.[1] The economy is endowed with L units of labor, which is the only factor of production. All of this labor is inelastically supplied at the wage, w_τ, where τ indexes time. Firms in the economy are small, in the sense that each firm produces only one variety, while taking the prevailing wage as given. (In the aggregate, however, w_τ will be an outcome of the general equilibrium of the model.)

Each unit of labor (or simply "worker") can engage in one of two activities, namely either in the production of differentiated varieties or in R&D activity. With regard to the former, production takes place under a constant returns-to-scale technology. Let $Z_\tau(j)$ denote the labor productivity associated with the best available idea for producing variety j at time τ; $1/Z_\tau(j)$ units of labor are thus required to produce each unit of this variety. Then, the unit cost faced by the firm that produces this variety is simply: $w_\tau/Z_\tau(j)$.

On the other hand, the objective of R&D activity is to generate ideas to improve upon existing technologies. Each idea spells out a technology (equivalently, a labor productivity level) for a specific differentiated variety. We model the generation of these ideas as a Poisson process with a constant arrival rate of λ for each R&D worker.[2] Following Kortum (1997) and Klette and Kortum (2004), conditional on receiving a new idea, the identity of the variety to which the idea applies is determined by a random draw from a uniform distribution on the unit interval.[3]

We specify a setting in which the innovation process is strictly cumulative. In particular, knowledge about production technologies diffuses immediately as soon as the good in question is marketed, so that the underlying knowhow becomes available to all agents in the economy. For example, this could be because it is easy to reverse-engineer the technology after observing a physical sample of the good. As the current best technology for producing each marketed good is widely-known, subsequent innovation effort strictly builds upon this knowledge to generate productivity improvements. In equilibrium, the best patented technology for each differentiated variety will indeed be used in production, with the good

[1] It is straightforward to extend the model to include a non-innovating outside sector, whose output can then play the role of the numeraire.

[2] In other words, the probability that an individual worker will receive a new idea during a small time interval $\Delta\tau$ is given by $\lambda\Delta\tau$. Moreover, each R&D worker can receive only one idea at any instant in time.

[3] As in these preceding papers, this rules out the possibility that innovation effort can be directed toward the production of specific varieties.

being marketed, and hence each subsequent arriving idea will always improve upon the frontier patented technology for the variety to which it applies.[4]

In the absence of IPR protection, the diffuse nature of knowledge would provide little incentive for private agents to undertake R&D. We therefore require that an IPR regime be in place that allows any new idea to be patented at negligible cost. By patenting, the firm in possession of the new idea gains exclusive rights to produce and market the variety (say, variety j) with the new technology, and will indeed have the entire market for j to itself as it is now the most productive manufacturer of j. This monopoly power only expires when the next idea that improves upon the technology for j arrives.[5] Ideas that are not patented but which are marketed can immediately be legally imitated by other firms, which would compete away the profits accruing to the innovator. It follows that firms will immediately patent any new ideas that they receive, so that no goods will be marketed without first being patented.

Having spelt out the arrival process for ideas, we now describe what governs the productivity levels associated with these ideas. To initialize the innovation process, we assume that at the start of time ($\tau = 0$), there is a baseline technology for each variety that is freely available to all firms. We normalize the productivity of this baseline technology to be 1 for all varieties, namely: $Z_0(j) = 1$ for all $j \in [0, 1]$. Now, define $Z^{(k)}(j)$ to be the productivity associated with the k-th idea to arrive (after time 0) for variety j, where k is a non-negative integer. Thus, $\{Z^{(0)}(j), Z^{(1)}(j), Z^{(2)}(j), \ldots\}$ form a sequence of the successive best technologies for producing this variety. To describe how this frontier technology evolves, define $\zeta^{(k+1)}(j) \equiv Z^{(k+1)}(j)/Z^{(k)}(j)$ to be the productivity improvement associated with the next idea to arrive. We specify $\zeta^{(k+1)}(j)$ to be a random variable that is an independent draw from the following standardized Pareto distribution with shape parameter $\theta > 1$:

$$\Pr\left(\zeta^{(k+1)}(j) < z\right) = 1 - z^{-\theta}, \quad \text{where } z \in [1, \infty), \text{ for all } k \geq 0. \tag{1}$$

Note that a lower θ implies a more fat-tailed distribution which places greater weight on drawing relatively large productivity improvements. For simplicity, the distribution in (1) does not depend on j, so that the underlying innovation process is symmetric across varieties. Moving forward, we will thus write $Z^{(k)}(j)$ simply as $Z^{(k)}$, since the distribution of the productivity level of the k-th idea to arrive will be identical for all varieties.[6]

The expression in (1) embodies the notion of cumulative innovation, since the lower bound of the support of the distribution of productivity improvements is 1. In effect, after the k-th idea has arrived, the productivity $Z^{(k)}$ associated with that idea becomes the new knowledge frontier which the $(k+1)$-th idea will improve upon. Note also that we have assumed that the distribution in (1) does not depend on how many ideas have already arrived (k) or on the productivity level of the last drawn idea ($Z^{(k)}$). In sum, this means that conditional on the realized value of $Z^{(k)}$, the next arriving idea $Z^{(k+1)}$ can be viewed as a draw from a Pareto distribution with the same shape parameter but with a lower bound of $Z^{(k)}$.[7]

At this juncture, it is useful to discuss the relationship between the innovation process that we have just described and that advanced in Kortum (1997) and Eaton and Kortum (2001). In the notation that we have adopted, the analogue of their specification for the (stationary) distribution that governs innovation is:

$$\Pr(Z^{(k+1)} < z) = 1 - z^{-\theta}, \quad \text{where } z \in [1, \infty), \text{ for all } k \geq 0. \tag{2}$$

[4] Alternatively, the innovation process can be set up as one entailing improvements along a quality dimension, where each arriving idea yields a higher utility to consumers with no change in the good's production cost (and hence market price).

[5] Given the continuous measure of varieties, there is a zero probability that the same agent will consecutively receive two ideas for producing the same variety.

[6] This Pareto specification for each productivity improvement is also adopted by Koléda (2004), Minniti et al. (2011), and Desmet and Rossi-Hansberg (2012). In particular, Minniti et al. (2011) provide descriptive evidence of: (i) substantial cross-firm heterogeneity in the usefulness of innovations (as captured by patent citations), and (ii) the Pareto providing a reasonable fit to the distribution of the value of patents especially in its right-tail.

[7] Recall that if a Pareto distribution is truncated from the left, the resulting distribution remains Pareto with the same shape parameter, but with the left truncation value serving as the new lower bound of its support.

Thus, in this earlier work, ideas that arrive may or may not surpass the current state-of-the-art technology, $Z^{(k)}$; those ideas that fall short of the frontier are not competitive enough to survive in the market. As more ideas accumulate over time in their economy, it becomes less likely that a new idea will surpass the current frontier. In contrast, our interest lies in understanding an innovation process in which each new idea strictly improves upon $Z^{(k)}$. The two approaches therefore represent two opposite ends of the spectrum: While Kortum (1997) and Eaton and Kortum (2001) adopt a non-cumulative formulation, we instead explore a situation where innovation is fully cumulative, this being motivated by our interest in analyzing the externalities that arise from R&D activity in this latter setting.

1.2 General equilibrium

We now embed the above cumulative innovation process in a general equilibrium setting.

Utility: The utility function of the representative consumer as of date 0 is given by:

$$U_0 = \int_0^\infty e^{-\rho\tau} \ln u_\tau \; d\tau. \tag{3}$$

Here, ρ is the rate of time preference (a parameter), while u_τ aggregates the instantaneous utility from the consumption of differentiated varieties at time τ. Specifically, u_τ is given by:

$$u_\tau = \exp\left\{\int_0^1 \ln x_\tau\left(j\right) \; dj\right\}, \tag{4}$$

where $x_\tau\left(j\right)$ denotes the quantity of variety j consumed at time τ.

The representative consumer chooses $\{x_\tau\left(j\right)\}_{\tau=0}^\infty$ in order to maximize (3), subject to the intertemporal budget constraint:

$$\int_0^\infty e^{-r\tau} X_\tau \; d\tau \leq b(0), \tag{5}$$

where $X_\tau = \int_0^1 p_\tau\left(j\right) x_\tau\left(j\right) dj$ denotes the flow of consumption spending at time τ, with $p_\tau\left(j\right)$ being the corresponding price of variety j at that time (which the consumer takes as given). r is the prevailing interest rate, which will be pinned down by the rate of return earned on owning a patent, this being the only asset in our economy. Finally, $b(0)$ denotes the present value of the future stream of wage income that will be earned by each consumer plus the value of her initial asset holdings at date 0.

It is well-known (see for example, Grossman and Helpman, 1991) that the solution to this dynamic optimization problem yields:

$$r = \rho + \frac{\dot{X}}{X}, \tag{6}$$

where $\dot{X}$ is the time derivative of X. (We will omit the time subscript for equations that hold for all $\tau \geq 0$.) Thus, the rate of growth of consumption spending should equal the difference between the market interest rate and one's private rate of time preference. It will now be convenient to set aggregate consumption expenditure $E_\tau \equiv LX_\tau$ as the numeraire for each τ. Since $E_\tau = 1$ and L is constant over time, it follows from (6) that $r = \rho$. Moreover, the expenditure on each variety will be constant and equal to 1 at each date, since we have a unit measure of varieties. As the expenditure on each j is invariant to its price, the price elasticity of demand for each variety is -1.

Market structure and profits: Firms compete by setting prices. If no ideas have yet arrived for variety j by time τ, then that variety is priced at marginal cost (w_τ), since the baseline technology is freely accessible to all potential producers.

On the other hand, if at least one idea has arrived for variety j, then firms compete under Bertrand competition. The firm possessing the most productive idea will set a limit price that is just enough to keep the second most productive firm (and by implication, all other firms) out of the market. Therefore, the equilibrium price for variety j at time τ when precisely k ideas have arrived is equal to: $p_\tau(j) =$

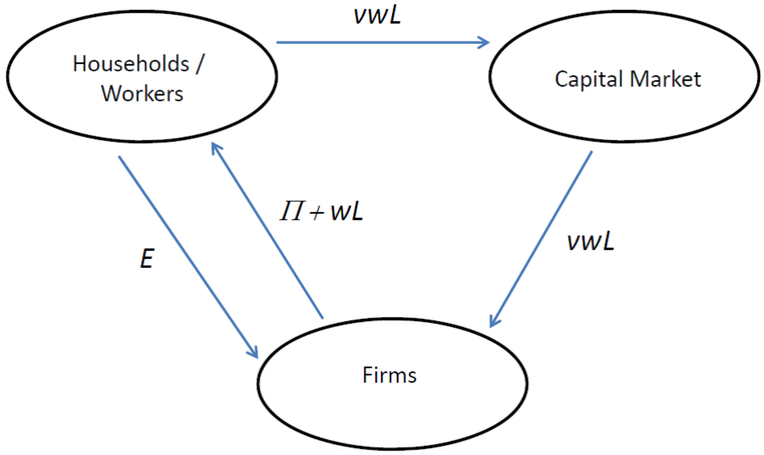

Figure 1: Circular Flow in the Economy

$w_\tau/Z^{k-1}(j)$. The price markup that the most productive firm sets is simply: $Z^k(j)/Z^{k-1}(j)$, so that it inherits the Pareto distribution from (1). To be more explicit, denoting $\mu(m)$ as the cdf of the price markup m, we have: $\mu(m) = 1 - m^{-\theta}$ for $m \geq 1$. The expected flow profits earned by a firm which holds the patent for the best technology for a given variety can now be computed as:

$$\Pi_\tau = \int_{m=1}^{\infty} \left(\frac{m-1}{m}\right) E_\tau \; d\mu(m) = \frac{1}{1+\theta}. \tag{7}$$

The above makes use of the fact that E_τ is also the expenditure per variety, since we have a unit measure of differentiated varieties.[8]

Savings and investment: Let $v \in [0,1]$ denote the share of the labor endowment L that is hired by firms to engage in R&D activity. (The remaining fraction, $1-v$, works in the production of varieties.) We assume that firms need to obtain financing in order to hire R&D workers. This spending on R&D is the only form of investment in our model, in the sense that the innovation is undertaken to generate ideas that yield a future stream of profits. In exchange for the financing they obtain, firms issue claims on the flow of profits from their patents. These claims (which we can think of as equity) are the only assets in the economy, and the total value of these assets at time τ is denoted by A_τ. (Note that A_τ is also equal to the value of each patent, given that we have a unit measure of varieties and only one active patent for each variety.) The above investment in R&D activity is financed through the savings of workers. This could take place with savers directly owning the equity of firms, or with the financing channeled through an intermediary such as a bank. Figure 1 summarizes this circular flow of funds in the economy between workers, firms and the financial intermediary sector (or capital market).

Bear in mind that workers have two sources of income, namely their labor income and the return on assets. The aggregate savings in the economy are thus equal to total income net of consumption spending, $wL + rA - E$. Equating this with aggregate net investment ($\dot{A}$) at each point in time, we have:

$$wL + rA - 1 = \dot{A}. \tag{8}$$

Research incentives: Since vL workers are employed in R&D, the Poisson arrival rate for new ideas in the economy as a whole is equal to $\lambda vL/\phi$, where ϕ is the difficulty of doing R&D. Moreover, we assume that

$$\frac{\dot{\phi}}{\phi} = \mu \frac{\lambda vL}{\phi} \quad \text{where } \mu \text{ is a constant}$$

[8] To be clear, Π is equal to profits for a variety conditional on at least one idea having arrived for the variety in question. In particular, Π is not equal to aggregate profits in the economy.

The rate of return r for owning an asset (namely, a patent) must equal the flow profit rate minus the probability of a complete capital loss due to the arrival of a new idea that supersedes the existing patent. This implies:

$$r = \frac{\Pi}{A} - \frac{\lambda v L}{\phi}. \tag{9}$$

As $r = \rho$, one can rewrite (9) as $A = \Pi/\left(\rho + \lambda v L/\phi\right)$, which gives us an expression for the expected present discounted value of each patent. Intuitively, this is equal to the present value of flow profits, discounted by the rate of time preference plus the hazard rate of losing the market to a subsequent innovator.

Labor market equilibrium: We consider an equilibrium in which a positive amount of production (and hence consumption) takes place in each time period. This implies that the wage of a production worker needs to weakly exceed the value of the marginal product of being an R&D worker. The latter is given by the flow rate of ideas that each R&D worker can generate multiplied by the value of each idea, namely $\lambda A/\phi$. We thus have:

$$\frac{\lambda A}{\phi} \leq w. \tag{10}$$

Equation (10) will hold with equality when some innovation activity takes place, namely when v lies in the interior of $[0, 1]$. Workers would then be indifferent between being employed in R&D and production.

Steady state: The five equations (6), (7), (8), (9) and (10) define a system in the five unknowns Π, A, w, r, and v, which pins down the steady state of our model.

In what follows, we focus on a steady state in which some innovation occurs, and in which the share of the labor force employed in R&D is constant over time. In particular, this means that (10) will hold as an equality. A quick inspection of our system of equations then implies that the value of a patent (A), the return on assets (r), and the return to labor (w) will all be constant in this steady state. Moreover, a familiar set of arguments can be applied to show that this steady state is one to which the economy immediately jumps. To see this, (8) and (10) together imply that: $\dot{A}/A = \lambda L/\phi + \rho - (1/A)$. If $\lambda L/\phi + \rho$ were to exceed $1/A$ at any time along the transition path, $\dot{A}/A$ would be positive, and the subsequent increase in A would further widen the gap between $\lambda L/\phi + \rho$ and $1/A$ on the right-hand side. Also, the larger is A, the faster is the rate of increase in A, so that the value of a patent would continue to increase to infinity. However, (9) places an upper bound of Π/ρ on the value of a patent, so that the expectation that A will increase indefinitely cannot be met. A similar argument can be used to rule out the reverse case where $\lambda L/\phi + \rho$ falls short of $1/A$ on the transition path. Thus, expectations about the value of a patent can only be fulfilled if the economy jumps immediately to a situation where $\dot{A}/A = 0$.

It is now straightforward to solve the system of five equations after setting $\dot{X} = \dot{A} = 0$. This yields in particular the following expression for the market allocation of labor to R&D activities, v^{eqm}:

$$v^{eqm} = \frac{\frac{\lambda L}{\phi} - \rho\theta}{\frac{\lambda L}{\phi}(1+\theta)} \tag{11}$$

Note from the above that v^{eqm} is clearly less than 1. To further ensure that $v^{eqm} > 0$, we need to impose the following:

Assumption 1: $\lambda L/\phi > \rho\theta$.

Intuitively, for there to be a positive amount of R&D in the steady state, we require that: (i) the innovative capacity of the economy (captured by $\lambda L/\phi$) be sufficiently high; (ii) the dispersion of the ideas distribution be large (θ small); and/or (iii) consumers be sufficiently patient (ρ low).

Using (11), one can verify directly that the research intensity of the economy varies naturally with the underlying parameters of the model. Firms hire a greater share of the workforce in R&D when the arrival rate of ideas is higher ($dv^{eqm}/d\lambda > 0$), or when those ideas are drawn from a Pareto distribution with a fatter right-tail ($dv^{eqm}/d\theta < 0$). Moreover, if agents are more patient when valuing future relative to current consumption, this also raises R&D effort in the steady state ($dv^{eqm}/d\rho < 0$).

1.3 Welfare

We turn next to the task of evaluating country welfare, in order to facilitate our later analysis of the efficacy of patent policy. The utility specification in (3) and (4) implies that welfare depends on the real wage in each period, as: $u_\tau = w_\tau / \exp\left\{\int_0^1 \ln p_\tau(j) dj\right\}$. (Recall in particular that the economy jumps immediately to its steady state.) Since all varieties are *ex ante* symmetric and we have a unit measure of these varieties, the law of large numbers implies that the ideal price index in the denominator is equal to: $\exp\{E[\ln P_\tau]\}$, where P_τ is a random variable whose realization is the price of a variety at time τ; the expectation operator is taken over this price distribution.

We therefore need to understand how prices evolve over time. Due to the Poisson nature of the innovation process, the probability that exactly k ideas have arrived by time τ when vL units of labor are engaged in R&D at each date is: $\frac{(\lambda v L\tau/\phi)^k}{k!} e^{-\lambda v L \tau/\phi}$, where k is a non-negative integer. Recall that when $k = 0$, the variety in question will be priced at w_τ (its marginal cost). On the other hand, when $k \geq 1$, under the limit-pricing rule, the price of a variety will instead be a random variable that inherits the distribution of $w_\tau / Z^{(k-1)}$. The expected log price of a variety at time τ is thus:

$$E\left[\ln P_\tau\right] = \frac{\left(\lambda v L\tau/\phi\right)^0}{0!} e^{-\lambda v L\tau/\phi} \ln w_\tau + \sum_{k=1}^{\infty} \frac{\left(\lambda v L\tau/\phi\right)^k}{k!} e^{-\lambda v L\tau/\phi} \left(\ln w_\tau - E[\ln Z^{(k-1)}]\right). \qquad (12)$$

Note that the first term and each term in the summation in (12) is equal to the probability that k ideas have arrived between times 0 and τ, multiplied by the log price at time τ when there have indeed been exactly k ideas (where $k = 0, 1, 2, ..., \infty$).

We show in the Appendix how to evaluate (12) explicitly. The key to this is to recognize that in the underlying innovation process, the random variable $Z^{(k-1)} = Z^{(k-1)}/Z^{(0)}$ is the product of $k-1$ independent realizations from the standardized Pareto distribution given earlier in (1). (Recall that $Z^{(0)} = 1$.) Building off this observation, one can show that $\ln Z^{(k-1)}$ is a random variable from a Gamma distribution with mean $E[\ln Z^{(k-1)}] = (k-1)/\theta$ (see the Appendix).[9] The expected log productivity of the k-th idea to arrive thus increases linearly in k, while increasing also in the thickness of the right-tail of the Pareto distribution from which the productivity improvements are drawn. Substituting this expression for $E[\ln Z^{(k-1)}]$ into (12) and simplifying, one then obtains: $E\left[\ln P_\tau\right] = \ln w_\tau + \frac{1}{\theta}\left(1 - \lambda v L\tau/\phi - e^{-\lambda v L\tau/\phi}\right)$.[10]

It follows that per-period utility (the real wage) is given by: $u_\tau = \exp\left\{-\frac{1}{\theta}\left(1 - \lambda v L\tau/\phi - e^{-\lambda v L\tau/\phi}\right)\right\}$.[11] Defining the growth rate of the real wage to be $g_\tau \equiv d\ln\left(u_\tau\right)/d\tau$, we have:

$$g_\tau = \frac{\lambda v L}{\theta\phi}\left(1 - e^{-\lambda v L\tau/\phi}\right), \qquad (13)$$

which is clearly positive when $v > 0$. Although the economy jumps immediately to a steady state in which A, Π, and w (the nominal wage) are constant, the real wage nevertheless rises over time as varieties are on average becoming cheaper when there is a positive amount of R&D. In other words, Assumption 1 which guarantees that $v^{eqm} > 0$ also ensures that $g_\tau > 0$ for all $\tau \geq 0$. Substituting in the expression for v^{eqm} from (11), one can further verify that: $dg_\tau/d\lambda > 0$ and $dg_\tau/d\theta < 0$. Thus, a higher arrival rate of

[9] To be absolutely precise, this statement about the distribution of $\ln Z^{(k-1)}$ holds only for $k \geq 2$. Nevertheless, when $k = 1$, we have that $E[\ln Z^{(0)}] = 0$, so that the formula $E[\ln Z^{(k-1)}] = (k-1)/\theta$ is also valid for $k = 1$.

[10] Much work has been done documenting the fit of the Pareto distribution for firm size distributions (e.g., Axtell 2001; Luttmer 2007; Arkolakis 2011). Interestingly, the Gamma distribution also features a thick right-tail, although it matches the empirical distribution of US firms less well for the largest firm sizes (Luttmer 2007).

[11] The nominal wage can also be solved for explicitly from the system of five equations that pin down the steady state. This is given by: $w_\tau = \lambda/(\rho + \lambda L/\phi)$.

ideas (higher λ) and a larger average productivity improvement (smaller θ) both raise the growth rate of the real wage at each date τ. From (13), one can moreover see that the growth rate of the real wage rises over time ($dg_\tau/d\tau > 0$): From an initial value of $g_\tau = 0$, this asymptotes toward a maximum growth rate of $\lambda vL/(\phi\theta)$. This property derives from the fact that as time progresses, the baseline technology is shed from use for a greater and greater share of varieties. As the first idea arrives for successive varieties, the innovation process gets jump-started for a greater measure of varieties in the unit interval, hence causing the overall growth rate to rise over time. However, this effect peters out, as the first idea eventually arrives in expectation for all varieties.

It is instructive here to compare the above against the properties of the models in Kortum (1997) and Eaton and Kortum (2001), which also focus on a steady state in which the share of the workforce employed in R&D is constant. In these preceding papers, innovation is not cumulative in nature, and perpetual growth in real wages is sustained instead by a growing R&D workforce (vL), which grows at the same exogenous rate as the labor force (L). Thus, more ideas are drawn in each period by the ever-growing number of R&D workers, overcoming the fact that it gets harder and harder for each idea drawn from the stationary distribution in (2) to surpass the technological frontier. In contrast, the model which we have just presented generates steady-state growth in real wages through the cumulative nature of innovation – new ideas always strictly improve on the technological frontier – without requiring that the labor force grow over time.

Finally, the expected welfare of the representative consumer is obtained by substituting the expression for u_τ into (3) and evaluating the associated integral. After some algebraic simplification, this yields:

$$U_0 = \frac{(\lambda vL/\phi)^2}{\rho^2\theta(\rho + \lambda vL/\phi)} = \frac{(\lambda L/\phi - \rho\theta)^2}{\rho^2\theta(1+\theta)(\rho + \lambda L/\phi)}, \tag{14}$$

where the last equality follows from replacing v by the expression for v^{eqm} from (11). One can show via straightforward differentiation that so long as Assumption 1 holds, (14) is increasing in λ and decreasing in θ. Welfare therefore rises either as innovations arrive more frequently or as the average productivity improvement increases.

1.4 Contrast with the social optimum

To understand the efficiency properties of the steady state which we have just solved for, it is instructive to compare the above market equilibrium with the outcomes under a benign social planner. Conceptually, this social planner's problem can be formulated as a labor allocation decision over the share of labor to employ in R&D, as well as the value of $L^p_\tau(j)$ for each $j \in [0,1]$, namely the amount of labor assigned to the production of variety j at each point in time. Formally, the social planner sets out to solve:

$$\max_{v,\{L^p_\tau(j)\}_{j=0}^1} \quad U_0$$

$$\text{s.t.} \quad \int_0^1 L^p_\tau(j)\, dj = L(1-v) \quad \text{for all } \tau \geq 0, \tag{15}$$

$$\text{and} \quad Lx_\tau(j) = L^p_\tau(j) Z_\tau(j) \quad \text{for all } j \in [0,1] \text{ and } \tau \geq 0. \tag{16}$$

The first constraint (15) is a labor market-clearing condition that states that all labor not engaged in R&D must be employed in production. On the other hand, the second constraint (16) sets the quantity demanded of variety j equal to the quantity produced at each period in time.

As we show in the Appendix, the solution to this social planner's problem features an equal allocation of labor to the production of each variety. In other words, given the choice of v, we have: $L^p_\tau(j) = L(1-v)$. Using this property, the planner's problem can then be simplified to the following unconstrained maximization problem over v:

$$\max_v \; U_0 = \frac{\ln(1-v)}{\rho} + \frac{\lambda vL}{\rho^2\theta\phi}.$$

The above maximand is a concave function in v and thus yields a unique optimal allocation of labor between research and production activities. This social planner's allocation, denoted by v^{SP}, is given by:

$$v^{SP} = \frac{\lambda L/\phi - \rho\theta}{\lambda L/\phi}. \tag{17}$$

This lies strictly in the interior of $[0,1]$ if $\lambda L/\phi > \rho\theta$, namely if Assumption 1 holds. Moreover, comparing this with the allocation that would emerge in the market equilibrium from (11), one immediately has the following result:

Proposition 1 *The share of labor that a social planner would allocate to research is strictly larger than that which is observed in the market equilibrium, namely* $v^{SP} > v^{eqm}$.

The decentralized equilibrium in our model therefore unambiguously yields less investment in R&D effort relative to the socially-optimal level. One can moreover see that the relative extent to which v^{eqm} falls short of v^{SP}, namely $(v^{SP} - v^{eqm})/v^{SP}$ is increasing in θ. Intuitively, the less fat-tailed is the Pareto distribution of productivity improvement draws, the less attractive are the potential private returns (profits) from R&D, and hence the greater the extent of under-investment in R&D in the market equilibrium relative to the social optimum.

The literature on endogenous growth in the presence of knowledge spillovers has highlighted several externalities that drive a wedge between the market and social-planner outcomes (e.g., Grossman and Helpman, 1991; Aghion and Howitt, 1992), and these forces are present too in our model. First, there is an "intertemporal spillover" effect arising from the cumulative nature of innovation: Firms apply a higher effective discount rate when evaluating the value of a patent because they do not internalize the positive knowledge spillovers from their innovation on future productivity improvements. Second, there is an "appropriability" effect, in that the private profits which firms earn are in general smaller than the full gains to consumer surplus that each innovation generates. Third, a "business-stealing" effect is at play, since innovation effort erodes the profits of preceding innovators in a way that a social-planner would want to fully internalize. The first two of these effects tend to decrease R&D in the market equilibrium relative to the planner's problem, while the last effect pushes firms toward over-investing in R&D. Proposition 1 implies that in our model, the former two effects must dominate the latter "business-stealing" mechanism.[12]

We can in fact make a more precise statement concerning the relative importance of these three externalities. By rearranging (11), observe that the market allocation of labor to research activity is determined as the solution to: $\frac{\lambda(1-v)L/\phi}{(\rho+\lambda v L/\phi)\theta} = 1$. On the other hand, the first-order condition of the social planner's problem implies that v^{SP} solves: $\frac{\lambda(1-v)L/\phi}{\rho\theta} = 1$. Thus, the only wedge between the two solutions arises from the different discount rates that are respectively applied: In the market equilibrium, firms use a higher discount rate of $\rho + \lambda v L/\phi$, which takes into account the flow probability of suffering a complete profit loss to a new innovation, on top of the social discount rate. The only externality that is relevant in our model is thus the intertemporal spillover effect; evidently, the appropriability and business-stealing effects must offset each other exactly.

1.5 Patent breadth policy and equilibrium research effort

With the above formulation of the patenting process, we can readily embed our model in a general equilibrium setting following the approach in Section 2.2. As explained above, we consider a situation

[12] Note that the presence of monopoly-pricing power *per se* does not distort labor allocations in our model. The reason is that all firms charge the same markup in expectation (drawn from the standardized Pareto distribution, $\mu(m)$), so that the allocation of production labor across varieties cannot be improved upon *ex ante*. See the related discussion in Grossman and Helpman (1991), p.70.

in which the NIS requirement, B, is introduced at date 0 and held constant subsequently. We highlight below how this policy intervention alters profits and research incentives relative to the baseline model.

We describe first the prices that will be observed. For a given variety j, if no patentable ideas have arrived by time τ, firms will produce this variety using the publicly-available baseline technology and price the good at marginal cost, w_τ. On the other hand, if $k \geq 1$ patentable ideas have arrived for variety j by time τ, Bertrand competition then implies that the firm with the best patentable idea will set a limit-price equal to: $w_\tau / \tilde{Z}^{(k-1)}(j)$. Once patentable ideas have started arriving, the firm with the best patented idea will price at a markup given by: $\tilde{Z}^{(k)}(j)/\tilde{Z}^{(k-1)}(j)$, which from (**??**) is a random variable with cdf: $\tilde{\mu}(m) = 1 - (m/B)^{-\theta}$, where $m \in [B, \infty)$ is the markup. The flow of profits accruing to this patent-holding firm can now be evaluated as:

$$\Pi_\tau = \int_{m=B}^{\infty} \left(\frac{m-1}{m}\right) E_\tau \, d\tilde{\mu}(m) = \frac{B(1+\theta)-\theta}{B(1+\theta)}. \tag{18}$$

(Recall that $E_\tau = 1$ by our choice of numeraire.) The above profit expression coincides with equation (7) from the baseline model when B is equal to 1. From (18), one can see that profits will take up a larger share of consumer expenditures when B is higher because patentable ideas embody a larger productivity improvement on average. As before, a smaller θ is also associated with higher profits, as a more fat-tailed Pareto distribution means that firms can on average expect to charge a higher markup.

The patenting requirement that is now in place will affect research incentives. As before, the value of assets in the economy is equal to the aggregate value of patents. The rate of return r to owning a unit of these assets must once again be equal to the profit rate net of the flow probability of experiencing a complete capital loss. On the latter, even though the Poisson arrival rate of ideas remains $\lambda v L/\phi$, each new idea will only be patentable with probability $B^{-\theta}$. In particular, a binding inventive step requirement ($B > 1$) would strictly lower the likelihood that an incumbent patent-holder gets superseded by a newly-arrived idea, which effectively extends the duration of the incumbent's monopoly power. At each point in time, we thus have:

$$r = \frac{\Pi}{A} - \frac{\lambda v L B^{-\theta}}{\phi}. \tag{19}$$

It follows from (19) that the expected present discounted value of each patent, A, is now equal to $\Pi / \left(\rho + \lambda v L B^{-\theta}/\phi\right)$.

Since the probability that a given idea will be patentable now depends on the required inventive step, the labor market equilibrium condition must also be modified accordingly. The value of the marginal product of an R&D worker is now $\lambda B^{-\theta} A/\phi$, as the flow probability of receiving an idea that clears the NIS requirement is $\lambda B^{-\theta}/\phi$. A more stringent requirement (a higher B) will thus reduce the expected returns to working in R&D. For some production to occur in each time period, we require that:

$$\lambda B^{-\theta} A/\phi \leq w, \tag{20}$$

namely that the wages from being a production worker weakly exceed the value of the marginal product of an R&D worker. In the steady state which we will consider, in which a positive and constant share of the workforce is allocated to research, the above must hold as an equality.

To close out the model, observe that the intertemporal welfare maximization and circular flow conditions from (6) and (8) remain unchanged. The equilibrium is therefore determined by these two equations in combination with (18), (19) and (20). The five unknowns of this system are once again Π, A, w, r, and v. As before, we focus our attention on a steady state in which v is (weakly) positive and constant. A set of arguments analogous to that in the baseline model can be applied to show that A, w and r will all be constant in this steady state. It will also be the case that the economy jumps instantaneously to this steady state upon the introduction of the inventive step requirement, B.

Equilibrium research effort: Setting $\dot{A} = \dot{X} = 0$ and solving out for this steady state, one obtains (after some algebraic simplification) the following expression for the R&D labor share:

$$v(B) = \max\left\{1 - \frac{\theta}{B(1+\theta)} - \frac{\rho\theta B^{\theta}}{\lambda L B(1+\theta)/\phi},\ 0\right\}. \tag{21}$$

We write v explicitly as a function of B to emphasize the scope that patent policy now has to affect equilibrium research effort. Note from (21) that $v(B)$ is strictly less than 1 for $B \geq 1$, so the steady state will never feature complete specialization in R&D. It is possible however for the steady state to feature no R&D, as $v(B) = 0$ when $B \to \infty$ (since $B^{\theta-1}$ would tend to ∞, as $\theta > 1$). To be precise, we have $v(B) > 0$ if and only if $\frac{\lambda L}{\phi} > \frac{\rho\theta B^{\theta}}{B(1+\theta)-\theta}$. When $B = 1$, this condition reduces back to Assumption 1 (which we will continue to adopt), and $v(B)$ coincides exactly with the expression for v^{eqm} from (11); in particular, this means that $v(1) > 0$. Straightforward differentiation further reveals that the function $\frac{B^{\theta}}{B(1+\theta)-\theta}$ is decreasing when $B \in [1, \theta^2/(\theta^2-1))$, and increasing when $B \in (\theta^2/(\theta^2-1), \infty)$, with the value of $\frac{B^{\theta}}{B(1+\theta)-\theta}$ tending to infinity as B gets arbitrarily large. This implies that there is a unique value of B, which we denote as B^0, below which $\frac{\lambda L}{\phi} > \frac{\rho\theta B^{\theta}}{B(1+\theta)-\theta}$, but above which the inequality will be violated. Thus, $v(B) > 0$ if and only if $B \in [1, B^0)$.

Holding B constant, various comparative static properties of $v(B)$ carry over from the baseline model. Specifically, research effort will be (weakly) higher if agents are more patient (ρ low), the average productivity improvement is larger (θ low), or the arrival rate of ideas is higher (λ large). More interestingly, we can characterize how patenting standards now affect R&D outcomes. From (21), the share of labor in research clearly varies non-monotonically with B. This is consistent with the observation that a more stringent inventive step requirement will have conflicting effects. On the one hand, a higher B lowers the hazard rate that an incumbent patent-holder faces of losing its market to a new innovation, which *ex ante* would raise incentives for firms to hire more R&D workers (a "profit" effect). However, a higher required inventive step also lowers one's probability of successfully obtaining a patentable idea in the first place, which is often termed the "hurdle" effect in the patenting literature.

To analyze the net effect of these two forces, let us define: $\tilde{v}(B) = 1 - \frac{\theta}{B(1+\theta)} - \frac{\rho\theta B^{\theta}}{\lambda L B(1+\theta)/\phi}$. We have:

$$\frac{d\tilde{v}}{dB} = \left(\frac{\theta}{1+\theta}\right)\frac{1}{B^2}\left[1 - \frac{\rho(\theta-1)}{\lambda L/\phi}B^{\theta}\right]. \tag{22}$$

Denote the value of B for which $d\tilde{v}/dB = 0$ by B^v, given explicitly by: $B^v = \left[\frac{\lambda L}{\rho(\theta-1)\phi}\right]^{\frac{1}{\theta}}$. Since $d\tilde{v}/dB < 0$ for all values of B lower than B^v, while $d\tilde{v}/dB > 0$ for all B above B^v, B^v is the unique turning point of $\tilde{v}(B)$. In addition, we have $B^v > 1$ so long as $\lambda L/\phi > \rho(\theta-1)$, which is automatically satisfied if Assumption 1 holds. Using these properties, we can now characterize the behavior of $v(B) = \max\{\tilde{v}(B), 0\}$. As illustrated in Figure 2, the equilibrium allocation of labor is positive at $B = 1$ and first rises as B is raised above 1. It then reaches its maximum value when $B = B^v$, before declining toward 0 and meeting the horizontal axis at $B = B^0$; we then have $v(B) = 0$ for all $B \geq B^0$.

Defining the rate of innovation as the Poisson arrival rate of ideas, $\lambda v L/\phi$, we now have:

Proposition 2 *Suppose that $\lambda L/\phi > \rho\theta$ (Assumption 1 holds). Then: (i) $\frac{dv}{dB} > 0$ when $B \in [1, B^v)$, (ii) $\frac{dv}{dB} < 0$ when $B \in (B^v, B^0)$, and (iii) $v(B) = 0$ for all $B \in [B^0, \infty)$, so that B^v is the unique value of the NIS requirement that maximizes the equilibrium allocation of labor to R&D. In particular, raising B slightly from $B = 1$ will induce a higher rate of innovation. However, raising B when the NIS parameter is already above B^v will lower the rate of innovation.*

Intuitively, when the inventive step requirement is smaller than B^v, the profit effect dominates the hurdle effect, so that a more stringent patent standard raises the incentive to employ labor in R&D.

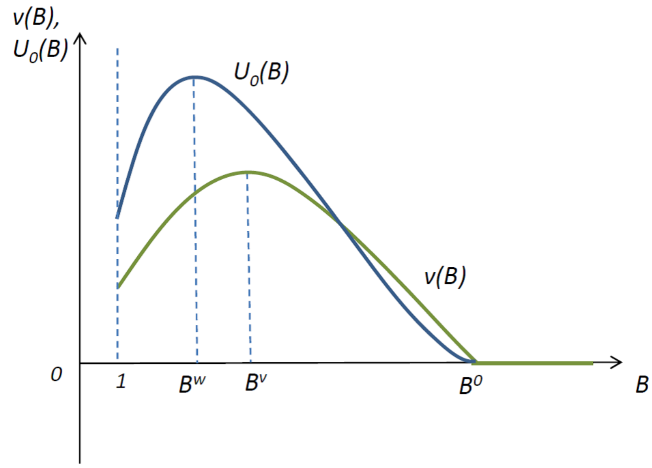

Figure 2: $v(B)$ and $U_0(B)$ (illustrated when $\lambda L > \rho\theta$)

However, when B exceeds B^v, the reverse holds, so that the strength of the hurdle effect now discourages research effort. In particular, this means that B cannot be set too high in practice, as research effort will eventually decline when the inventive step requirement is raised further. As highlighted in the Introduction, this provides one rationalization for the empirical observation in prior work of a weak and even negative association between the strength of patent protection and innovation outcomes, namely that the protection extended may be so generous to incumbent patent-holders as to suppress the aggregate research effort in the steady state.

An implication that emerges from the above discussion is that when $\rho(\theta-1) < \lambda L/\phi < \rho\theta$, the economy would feature no R&D effort when $B = 1$ (since $v(1) < 0$), but the incentive to undertake research would nevertheless be increasing in the NIS parameter when $B \in [1, B^v)$, since $dv/dB > 0$ in this range. If in addition $v(B^v)$ were strictly positive, setting $B = B^v$ would then tip the economy into a steady state with positive R&D effort. For an economy in these circumstances, the inventive step policy could then play a crucial role in inducing innovation and growth. The precise conditions under which this scenario arise are spelled out in the following:

Proposition 3 *Suppose that* $\rho(\theta-1)\left(\frac{\theta^2}{\theta^2-1}\right)^{\theta} < \lambda L/\phi < \rho\theta$. *There exists a range of binding inventive step policy parameters (with* $B > 1$*) that will shift the economy from a zero-growth to a positive-growth steady state. In particular, setting* $B = B^v$ *will achieve this shift.*

The conditions under which Proposition 3 hold describe an economy whose aggregate innovative capacity (as captured by $\lambda L/\phi$) lies in an intermediate range. This is the case when innovative capacity $\lambda L/\phi$ is low enough that no innovation arises when $B = 1$, but nevertheless high enough so that some government intervention can help to create research incentives. (The proof of this proposition is in the Appendix; we also show there that $(\theta-1)\left(\frac{\theta^2}{\theta^2-1}\right)^{\theta} < \theta$ is satisfied for all $\theta > 1$, so that the conditions for the proposition to hold can indeed be met.)

1.6 Patent breadth policy and welfare

We turn now to evaluate the consequences for welfare. As in our baseline model, this will require that we evaluate the ideal price index that consumers face, in order to compute their real wage. Note that the expected log price is now given by:

$$E\left[\ln P_\tau(j)\right] = \frac{\left(\lambda v L B^{-\theta}\tau/\phi\right)^0}{0!} e^{-\lambda v L B^{-\theta}\tau} \ln w_\tau + \sum_{k=1}^{\infty} \frac{\left(\lambda v L B^{-\theta}\tau/\phi\right)^k}{k!} e^{-\lambda v L B^{-\theta}\tau/\phi} \left(\ln w_\tau - E[\ln \tilde{Z}^{(k-1)}]\right). \tag{23}$$

The first term above is the expected log price for a variety when there are no patentable ideas, while the remaining terms in the summation are the corresponding expressions that apply when exactly $k \geq 1$ patentable ideas have arrived. There are two modifications in (23) relative to equation (12) from Section 1.3. First, the Poisson arrival rate is now $\lambda v L B^{-\theta}/\phi$, with the additional $B^{-\theta}$ term capturing the productivity hurdle that new ideas must cross; this tends to lower the arrival probability of patentable ideas. Second, when $k \geq 1$, firms now set their limit price at the marginal cost implied by the previous *patentable* idea, namely $w_\tau/\tilde{Z}^{(k-1)}$. Recall in particular that the productivity improvement between consecutive patentable ideas ($\tilde{Z}^{(k)}/\tilde{Z}^{(k-1)}$) is drawn from the Pareto distribution in (**??**) with the lower bound of its support equal to B.

Our analysis is once again tractable because the above price index can be worked out explicitly. We show in the Appendix that $\ln \tilde{Z}^{(k-1)}$ takes on the same Gamma distribution from the baseline model, but with a linear shift. Specifically, we find that: $E[\ln(\tilde{Z}^{k-1})] = (k-1)\left(\frac{1}{\theta} + \ln B\right)$ for all $k \geq 1$, with the additional $(k-1)\ln B$ term reflecting the effect that the inventive step policy has in raising the expected productivity level of the $(k-1)$-th innovation. Substituting this property into (23), one can then show that: $E\left[\ln P_\tau(j)\right] = \ln w_\tau + \left(\frac{1}{\theta} + \ln B\right)\left(1 - \lambda v L B^{-\theta}\tau/\phi - e^{-\lambda v L B^{-\theta}\tau/\phi}\right)$.

It follows that the real wage at time τ, which is also equal to the period flow utility u_τ, is given by: $w_\tau/E\left[\ln P_\tau(j)\right] = \exp\left\{-\left(\frac{1}{\theta} + \ln B\right)\left(1 - \lambda v L B^{-\theta}\tau/\phi - e^{-\lambda v L B^{-\theta}\tau/\phi}\right)\right\}$. At time τ, the real wage therefore grows at the following positive rate:

$$g_\tau \equiv \frac{d \ln u_\tau}{d\tau} = \left(\frac{1}{\theta} + \ln B\right) \frac{\lambda v L B^{-\theta}}{\phi} \left(1 - e^{-\lambda v L B^{-\theta}\tau/\phi}\right). \tag{24}$$

Given B, and hence $v(B)$, the growth rate of the real wage rises and asymptotes over time to a maximum of $\left(\frac{1}{\theta} + \ln B\right)\lambda v L B^{-\theta}$, for much the same reasons as were discussed in the baseline model.

The welfare of the representative consumer is then given by plugging in the above expression for u_τ into (3) and evaluating the integral for the present discounted value of the flow of real wages. This yields:

$$U_0(B) = \left(\frac{1}{\theta} + \ln B\right) \frac{\left(\lambda v L B^{-\theta}/\phi\right)^2}{\rho^2\left(\rho + \lambda v L B^{-\theta}/\phi\right)}. \tag{25}$$

Further replacing v in the above by the expression for $v(B)$ from (21), one then obtains a welfare formula that depends only on the primitive parameters of the model (ρ, θ, and $\lambda L/\phi$) and the required inventive step, B.

We can now assess the tradeoffs that arise from the use of this NIS requirement. Note from (25) that $U_0(B) = 0$ for all $B \geq B^0$, as $v(B) = 0$ in this range of values of B. We thus restrict our attention to study how welfare behaves when $B \in [1, B^0)$, where both $U_0(B)$ and $v(B)$ take on positive values. Differentiating the welfare expression in (25) with respect to B, one obtains:

$$\frac{dU_0}{dB} \propto \left(\frac{1}{\theta} + \ln B\right) B \frac{dv}{dB} - \theta v \ln B - \frac{\rho v}{2\rho + \lambda v L B^{-\theta}}, \tag{26}$$

where '$\propto$' indicates equality up to a positive multiplicative term. The first-order necessary condition for a local welfare maximum thus entails setting the right-hand side of (26) equal to 0.

Is the welfare-maximizing inventive step requirement binding (i.e., strictly greater than 1)? And does the welfare function in (25) indeed exhibit a unique maximum turning point? The former issue can be addressed by examining the behavior of dU_0/dB in the neighborhood of $B = 1$. One can verify through direct substitution that $dU_0/dB > 0$ at $B = 1$, as long as Assumption 1 holds, so that there is in fact a net gain from increasing the inventive step parameter slightly above 1. On the latter issue, even though $U_0(B)$ is in general not concave for all $B \in [1, B^0)$, we nevertheless can prove that the right-hand side of (26) is strictly decreasing in B in the smaller interval $B \in [1, B^v)$. (See the Appendix for details.) Observe too that for $B > B^v$, we have $dv/dB \leq 0$; from (26), this implies that $dU_0/dB < 0$ for $B > B^v$. Taken together, we find that dU_0/dB is first positive at $B = 1$, is strictly negative at $B = B^v$, and exhibits at most one root in the interval $[1, B^v)$. This allows us to conclude that there is a unique B^w that satisfies $dU_0/dB = 0$. Figure 2 illustrates these properties of the welfare function $U_0(B)$, in conjunction with the behavior of $v(B)$.

This leads us to our main result characterizing the optimal NIS requirement from a welfare perspective:

Proposition 4 *Suppose that $\lambda L/\phi > \rho\theta$ (Assumption 1 holds). Then the welfare-maximizing inventive step requirement is unique with $1 < B^w$.*

There is thus scope to improve welfare by setting an inventive step requirement strictly larger than 1, so long as λL is sufficiently large. The role of Assumption 1 here is intuitive: The innovative capacity of the economy needs to be high enough to ensure that the increased rate of innovation will more than exceed the social cost of ceding more monopoly power to patent-holders.

It is helpful at this juncture to examine (26) closely in order to get more intuition on the economic tradeoffs involved in the setting of the NIS requirement. The net effect of stronger patent protection is in principle ambiguous, but the underlying effects can be decomposed systematically. Note first that welfare can be written more explicitly as: $U_0 = U_0(B, v(B))$, with its corresponding total derivative given by: $\frac{dU_0}{dB} = \frac{\partial U_0}{\partial v}\frac{dv}{dB} + \frac{\partial U_0}{\partial B}$. The first term on the right-hand side of (26), namely $\left(\frac{1}{\theta} + \ln B\right) B \frac{dv}{dB}$, corresponds precisely to the $\frac{\partial U_0}{\partial v}\frac{dv}{dB}$ term in this total derivative. The key force captured here is commonly referred to in the IPR literature as the "dynamic" effect of patent protection in raising innovation rates (see for example, Nordhaus 1969; Tirole 1988; Grossman and Lai 2004). In the context of our model, we have already seen that when B is sufficiently small, specifically when $B \in [1, B^v)$, an increase in the required inventive step raises the steady-state allocation of labor to research ($\frac{dv}{dB} > 0$), thus raising the equilibrium rate of innovation. *Ceteris paribus*, this has a positive effect on welfare as $\frac{\partial U_0}{\partial v} > 0$, a fact which can be verified by straightforward differentiation of (25). Therefore, this dynamic effect term, $\frac{\partial U_0}{\partial v}\frac{dv}{dB}$, is indeed positive when $B \in [1, B^v)$.

This potential benefit from raising the NIS requirement needs to be weighed against a countervailing force, namely the static loss suffered in each period by consumers arising from the longer duration of each patent-holder's monopoly power. This latter effect is reflected by the second and third terms in (26), $-\theta v \ln B - \frac{\rho v}{2\rho + \lambda v L B^{-\theta}/\phi}$, which are clearly negative when $v > 0$. Note that these terms indeed map precisely to the $\frac{\partial U_0}{\partial B}$ term in the preceding total derivative. The welfare-maximizing inventive step requirement therefore needs to trade off the dynamic gains from a greater degree of patent protection against the static consumer surplus losses that are incurred.[13] When $B < B^w$, the dynamic effect dominates the static effect, and so $dU_0/dB > 0$ in this range of values of B; conversely, when $B > B^w$, the static effect dominates and $dU_0/dB < 0$.

Inspecting $\frac{dU_0}{dB} = \frac{\partial U_0}{\partial v}\frac{dv}{dB} + \frac{\partial U_0}{\partial B}$ further, we have: $dU_0/dB < 0$ for all $B \in [B^v, B^0)$. This holds because $dv/dB \leq 0$ when $B \in [B^v, B^0)$, and also because it is always true that $\frac{\partial U_0}{\partial B} < 0$ and $\frac{\partial U_0}{\partial v} > 0$. Welfare is therefore strictly decreasing in the interval $[B^v, B^0)$. Since the inventive step parameter that maximizes welfare is unique, B^w must lie in the interval $B \in [1, B^v)$ where $dv/dB > 0$, so that there is some benefit from introducing an inventive step policy through the increased research effort it induces. We sum up this argument as:

[13] A similar tradeoff is encountered if one considers instead the problem of choosing B to maximize the steady-state growth rate in (24). As discussed in section 3.2, the marginal benefit to innovators from raising B (profit effect) needs to be weighed against the marginal cost (hurdle effect).

Proposition 5 *The value of the required inventive step B that maximizes welfare is strictly less than that which maximizes the rate of innovation. In other words, $B^w < B^v$.*

This result is actually very intuitive, as the value of B that maximizes the research intensity of the economy requires consumers to give up too much current consumption to invest in R&D. When it is welfare instead that is the policy-maker's objective, an additional cost associated with raising B must be taken into account, namely the loss to consumer surplus. This insight is in fact a relatively general one: Horowitz and Lai (1996) also obtained an analogous result in a different setting where the policy instrument instead takes the form of a specified patent length, namely that the patent duration that maximizes the rate of innovation is longer than that which maximizes welfare.

1.7 Properties of the optimal patent breadth

Having established the existence of a unique welfare-maximizing NIS requirement, we turn next to explore how B^w is influenced by the deep parameters of our model. As we do not have a closed-form expression for B^w, we pursue a numerical approach to illustrate the behavior of this optimal inventive step requirement.

A convenient feature of our model is that the equilibrium is in fact characterized by a parsimonious set of three parameters. These are the discount rate (ρ), the shape parameter of the Pareto distribution (θ), and the innovative capacity of the economy ($\lambda L/\phi$). In particular, all steady-state outcomes depend only on the product $\lambda L/\phi$, and not on the separate values taken on by λ, ϕ and L. We proceed by adopting a set of baseline values for these three parameters, from which we illustrate the effects of varying each key parameter in turn. We should stress that the purpose here is not to offer a strict calibration, but rather to explore how B^w (and B^v) behave around a sensible choice of parameters.

We first choose $\rho = 0.07$. This follows Kortum (1997), who matches this to the real return observed on stock markets, which arguably provides a relevant reference point for the returns to innovation. That said, we will explore a wide range of values for ρ ranging between 0.02 and 0.12. For θ, we set this equal to 4; based on (7), this corresponds to expected profits making up a 25% share of expenditures per variety in the absence of a binding inventive step policy (as $1/(1+\theta) = 0.25$). We allow θ to vary between 2 and 6 in our numerical exercises, which translates into an expected profit share ranging from 14% to 33%. Last but not least, we pick $\lambda L/\phi = 1$, which implies a relatively modest average arrival rate of one new idea per variety in a given year. Note that this set of baseline parameters satisfies $\lambda L/\phi > \rho\theta$, which is required for our propositions to hold. They moreover imply a baseline welfare-maximizing NIS requirement of $B^w = 1.14$, with the steady-state share of labor allocated to R&D being $v(B^w) = 0.22$.[14]

The left-hand column of Figure 3 shows how the optimal inventive step requirement responds to the underlying primitives of the model. Observe first that a higher discount rate is associated with a lower B^w. Intuitively, as agents place a higher weight on current relative to future consumption, the marginal social benefit of promoting innovation is lower, and so the welfare-maximizing patent policy responds by placing less emphasis on promoting research. Not surprisingly, B^w is also lower if the idea distribution exhibits a thinner right-tail (higher θ), due to the smaller marginal benefit that can be gained from raising B when the productivity improvement draws are on average smaller. A similar logic explains why one should expect to see a higher optimal inventive step requirement in economies with a larger innovative capacity ($\lambda L/\phi$). Interestingly, among these three parameters, B^w appears most sensitive to changes in θ. The plots in the right-hand column of Figure 3 moreover confirm that B^v (the inventive step requirement that maximizes the allocation of labor to research) exhibits similar comparative statics. Note from the figures that B^v is always less than B^w, consistent with Proposition 5.

References

Paul S. Segerstrom, "Endogenous Growth without Scale Effects." The American Economic Review, Vol. 88, No. 5 (Dec., 1998), pp. 1290-1310.

[14] From (24), these parameter values also imply a maximum growth rate of the real wage of 4.8%, which is achieved asymptotically as $\tau \longrightarrow \infty$.

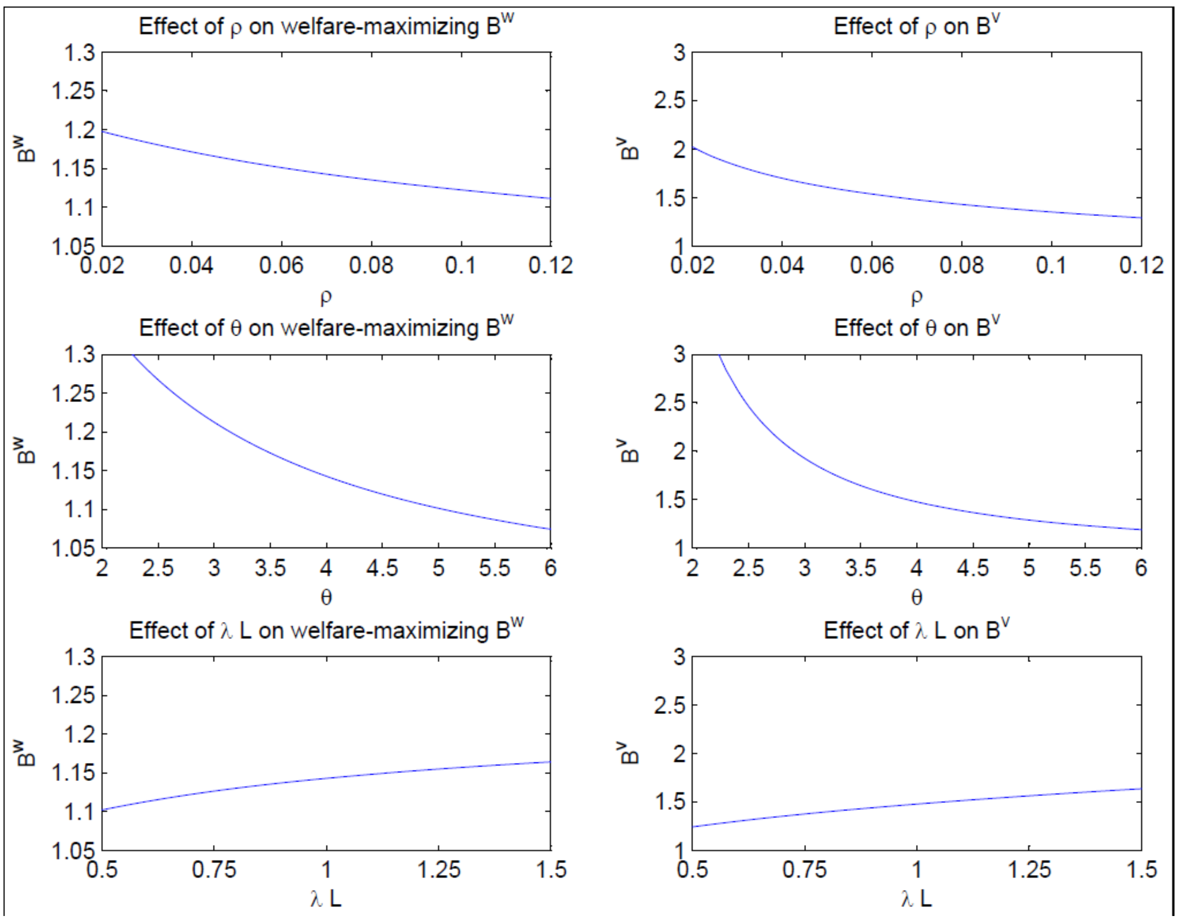

Figure 3: How B^w and B^v vary with parameter values